

MAJA N. VOLODINA (Hrsg.):

# Mediensprache und Medienkommunikation

im interdisziplinären und  
interkulturellen Vergleich

Mit einem einleitenden Beitrag  
von Ludwig M. Eichinger

INSTITUT FÜR DEUTSCHE SPRACHE



Werner Holly

## **Sprache in Massenmedien: Medialität, Intermedialität und kulturelle Praxis**

### **1. Für die Entwicklung einer kulturell fundierten Medienlinguistik**

Jedes Textstück aus einer Zeitung, jeder Ausschnitt aus einem Nachrichtenfilm oder einer Talkshow kann belegen, dass Sprachgebrauch nicht mehr mit dem Handwerkszeug traditioneller Sprachwissenschaft allein beschrieben und erklärt werden kann. Die vielfältigen Phänomene sprachlicher Medialität und ihre intermedialen Einbindungen erfordern eine Abkehr von reduktionistischen Sprachtheorien ebenso wie die Entwicklung einer Medienlinguistik, die mediale Texte als komplexe Bedeutungskonstrukte auf dem Hintergrund *technischer Dispositive* und *kultureller Praktiken* zu verstehen vermag.

### **2. Kleiner geschichtlicher Rückblick**

Schon die Differenz von gesprochener und geschriebener Sprache, von Mündlichkeit und Schriftlichkeit, lässt sich nicht ohne Berücksichtigung der technischen und kulturellen Rahmungen erfassen. Spätestens mit immer neuen Schreibgeräten und Trägermaterialien für Geschriebenes wird Technik für Sprachkommunikation konstitutiv. Ihre Beherrschung erfordert mühsam zu erlernende Kompetenzen, sie schafft Expertentum und differenziert bis heute Gesellschaften nach Graden der Bildung ihrer Mitglieder. Die Medientechnik entfaltet sich in immer neuen Kommunikationsformen (Papyri, Codices, Urkunden, Briefe, Bücher, Flugschriften, Zeitungen, Zeitschriften, Plakate), die auch die Formen ihrer institutionellen und sozialen Organisation einschließen: Kanzleien, Bibliotheken, Schreibmanufakturen, Universitäten, Druckereien, Verlage, Buchhändler, Redaktionen, Korrespondentennetze, Nachrichtenagenturen usw. Dabei vollziehen sich diese technischen und sozialen Entwicklungen als Formen kultureller (sprachlicher und kommunikativer) Praktiken, die in einem Wechselspiel sowohl als Verursacher als auch als Ergebnisse medialer Möglichkeiten zu sehen sind.

Mit den elektronischen Medien kommt seit dem ausgehenden 19. Jahrhundert auch die mündliche Kommunikation in den Sog der Technik. Mit Telefon und Schallplatte, Mikrofon, Lautsprecher, Radio, Tonband, Tonfilm, Fernsehen und Tonvideo wird auch die sprachliche Mündlichkeit Gegenstand einer rasanten

Entwicklung und wird von 'primärer' zu 'sekundärer Oralität' in immer neuen Ausprägungen, die mit der Spontansprache des Alltags oder der unverstärkten mündlichen Rede vor anwesendem Publikum nur scheinbar identisch sind. Zwar tendieren Medien dazu, sich selbst transparent zu machen und ihre Gemachtheiten zu naturalisieren, aber ihre Einflüsse sind auf allen linguistischen Ebenen (vom Laut zum Text) folgenreich, wie jede genauere Analyse zeigen kann. Der Triumphzug der audiovisuellen Medien Tonfilm und Fernsehen, die große Medienkonzerne haben entstehen lassen, zeigt, wie stark die Attraktion von Kommunikationsformen ist, die in der Lage sind, ähnlich wie primäre *face-to-face*-Begegnungen, jetzt aber unabhängig von Zeit- und Raumgrenzen, wieder Auge und Ohr in dynamischer Performanz einzubinden, und zwar scheinbar mühelos rezipierbar.

Die neueste Entwicklung, die durch die Stichworte Digitalisierung und Vernetzung zu kennzeichnen ist, hat nicht nur weitere Kommunikationsformen hervorgebracht (Elektronische Texte, Datenübertragung, Website, E-Mail, Chat, Foren, SMS, MMS), sondern auch neue Möglichkeiten ihrer Integration durch die gemeinsamen digitalen Formate, die auch alte Kommunikationsformen überarbeiten und einschließen können. Sprache und Schrift sind in einer neuen Weise manipulierbar und kommunizierbar, die noch kaum begriffen werden kann. Dabei ist aber jeweils darauf zu achten, dass nicht jede Art der Variation auf die mediale Kommunikationsform zurückgeführt werden kann. Dass dabei soziale Stile ins Spiel kommen, die mehr mit den Interessen und Gewohnheiten ihrer Nutzer zu tun haben als mit strikten medialen Faktoren, hat etwa die linguistische Beschreibung von E-Mails oder SMS gezeigt, die vorschnelle populärwissenschaftliche Identifizierungen von jugendsprachlichen Merkmalen (wie Inflektiven oder Emotikons) und Medien korrigieren musste.

Es kommt also darauf an, Medientexte sehr genau im Bedingungsgefüge jeweiliger Kommunikationsformen und Gattungen (wie z.B. 'internationale Nachrichtenagenturfilm' oder 'politische Talkshows') zu beschreiben: einerseits als *intermediale Zeichenkomplexe*, deren Materialität und Kommunikabilität nicht mehr nur von anthropologischen Gegebenheiten, sondern auch von deren technischen Überformungen ermöglicht und begrenzt sind; andererseits als *kulturelle Praktiken*, die von politischen und institutionellen Rahmungen ebenso wie von sozialen, ethischen und ökonomischen Regelhaftigkeiten und Interessen beeinflusst sind. Nicht alles ist dabei dem intentionalen Zugriff der Kommunizierenden zugänglich, manches geschieht gleichsam 'hinter dem Rücken' der Beteiligten, so dass der Eindruck einer 'Eigengesetzlichkeit' von Medien entsteht. Außerdem muss mit beträchtlichen Perspektivendivergenzen von Textproduzenten und Textrezipienten gerechnet werden.

Wie das Zusammenspiel von *Verfahren der Medialität* und *Mustern kultureller Praxis* zu massenmedialen Texten führt, die als multimodale und multi-kodale Komplexe der Bedeutungskonstitution verstanden werden können und müssen, soll nun an kleinen Beispielen der Audiovisualität von TV-Nachrichtenfilmen gezeigt werden.

### **3. Nachrichtenfilme als Beispiel: technisch-mediales Dispositiv und kulturelle Praktiken**

Nachrichtenfilme sind innerhalb von Fernsehnachrichtensendungen relativ kurze Elemente, die – oftmals in einen ganzen Block eingebettet – Bild- und Tonmaterial von internationalen Agenturen oder anderen Anbietern so montieren und mit einem Sprechertext unterlegen, dass sie zumindest eine minimale Information über ein aktuelles Ereignis zu liefern imstande sind. Es geht mir hier darum zu zeigen, dass audiovisuelle technische Medien ein bestimmtes semiotisches Potenzial haben, das im Zusammenspiel der möglichen Zeichenarten so genutzt werden kann, dass ihre jeweiligen semiotischen Schwächen nach Möglichkeit wechselseitig kompensiert werden. Dies geschieht auf der Basis der technisch-medialen Bedingungen der Kommunikationsform und in kulturell verankerten Praktiken.

#### **3.1 Das technisch-mediale Dispositiv: Nachrichtenfilmaudiovisualität**

Semiotisch sind Nachrichtenfilme deshalb aufschlussreich, weil hier das Bildmaterial eindeutig primär ist, der Sprechertext also nachträglich hinzugefügt wird, um die Informationsvermittlung zu sichern. Andererseits ist der Sprechertext nicht selten nach dem Vorbild von Zeitungsmeldungen zwar knapp, aber doch umfassend formuliert und damit quasi autark, so dass die Bilder nur noch illustrierenden Charakter zu haben scheinen. Im folgenden Beispiel des Sprechertextes eines Nachrichtenfilms des Zweiten Deutschen Fernsehens (ZDF) vom 1. Juli 2005 könnte man argumentieren, dass ein Linguist die Bild- und Tonebene nicht zu berücksichtigen braucht, um zu erklären, wie der Text der anonymen Sprecherin ('voice over') verstanden werden soll (ausführlicher in Holly 2011):

(I) in der türkischen hauptstadt Ankara hat die polizei offenbar einen bombenanschlag auf das justizministerium verhindert / (II) die beamten erschossen den mutmaßlichen attentäter nach einer kurzen verfolgungs-jagd auf offener strasse / (III) er soll zuvor versucht haben in das regierungsgebäude einzudringen und dort einen sprengsatz zu zünden / (IV) angeblich war er mitglied einer linksextremistischen terrorgruppe

Das ist im nächsten Beispiel (ausführlich dazu Holly 2009) etwas anders; hier ist der Sprechertext überhaupt nur nach einer vorausgegangenen Anmoderation verständlich, die Thema und Ort erläutert; aber auch dann reicht der Sprachtext allein nicht aus, man braucht die Bild- und Tonebene, um zu erkennen, dass ab (3) ein O-Ton-Sprecher eingebettet ist, der mit *wir* in (4) ganz anders referiert als die anonyme ‘Voice-over’-Sprecherin zuvor:

(1) noch wissen experten nicht wie viel der über zwei millionen liter öl des gestrandeten frachters bereits ausgelaufen sind (2) doch von flugzeugen aus ist bereits ein großer ölteppich zu beobachten eine bedrohung für viele meeressäuger fische und seevögel die in dem abgelegenen küstenabschnitt heimisch sind (3) hilfe ist kaum möglich (4) das gebiet ist sehr unwegsam und *wir* müssen ständig gegen die wellen ankämpfen [...]

Spätestens hier ist deutlich, dass die medialen und intermedialen Bezüge herangezogen werden müssen, um das Funktionieren dieses komplexen Textes zu verstehen, der bewegte Bilder, Töne und Sprachzeichen enthält. Dies gilt allerdings auch schon im ersten Beispiel, wo der Sprachtext zwar autark erscheint, aber wo es tatsächlich von Anfang an starke Sprach-Bild-Bezüge gibt. Dies zeigen schon die ersten beiden Einstellungen mit dem zugehörigen Sprachtext; in diesen beiden Einstellungen sieht man einen Mann aus einem Haus laufen, der über die Straße trabt, dann durch eine dünne Hecke in eine Parkstraße einbiegt, offensichtlich verfolgt von anderen, darunter Uniformierte, die mit Gewehren auf ihn zielen:



(1/I) in der türkischen hauptstadt Ankara hat die polizei offenbar einen bombenanschlag auf das justizministerium



(2) verhindert (II) die beamten erschossen den mutmaßlichen attentäter nach einer

Die folgende kurze Erläuterung soll verdeutlichen, wie in solchen Nachrichtenfilmen zwischen Sprach- und Bildebene Bezüge durch die bloße Montage hergestellt werden, so dass man unweigerlich versucht, im Bild zu sehen, was der Text sagt und umgekehrt im Text erklärt zu finden, was man im Bild sieht.

Am Beginn des Sprechertexts steht ein klassischer Lead-Satz, wie er gewöhnlich Nachrichten in Print- und Hörmedien eröffnet, ob Meldungen oder längere Berichte: (I) *in der türkischen hauptstadt Ankara hat die polizei offenbar einen bombenanschlag auf das justizministerium verhindert*. Er fasst das Wesentliche kurz zusammen und gibt einen Überblick, der den Ausgang der Geschichte vorwegnimmt. Hier überschreibt er die Einstellung (1, der laufende Mann), der man noch nicht entnehmen kann, was das Sichtbare mit dem Gesagten zu tun hat. Weder Schauplatz noch Akteure noch Ereignis sind visuell selbsterklärend, man nimmt aber auch schon visuell in der Schrifteinblendung den Ort des Geschehens („Ankara“) wahr. Wenn man die danach folgende Einstellung (2, die schießenden Uniformierten) sieht, ist man immerhin auf das Referenzobjekt ‘Polizei’ vorbereitet, das nun auch verständlich im Bild erscheint (man liest „Polis“ auf einer Jacke, man identifiziert nun Leute mit Gewehr als Polizisten); damit ist eine erste Brücke hergestellt. Die Bezugnahme vom Wort in (I) zum Bild in (2) ist gewissermaßen kataphorisch, vorausweisend, die vom Bild zum Wort anaphorisch, rückverweisend. Die Äußerung (II) greift das thematische Subjekt aus (I) (*die polizei*) substituierend auf (*die beamten*) und spezifiziert das Rhema von (I) (*bombenanschlag verhindert*) durch eine konkretisierende und personalisierende Formulierung (*erschossen den mutmaßlichen attentäter*), die nun auch das Anfangsbild lesbar macht, indem sie einen Einzelnen ins Spiel bringt, der als der anfangs Gezeigte erschossen werden kann.

Will man also die spezifischen Leistungen von Sprache und bewegten Bildern in diesem kurzen Zusammenspiel verstehen, so muss man sich auf die Aufgaben von Nachrichtenfilmen besinnen; es geht um die glaubwürdige und verständliche Information der Rezipienten. Man konnte aber schon an dem kurzen Ausschnitt sehen, dass die Bilder allein hier ein „Darstellungsproblem“ hätten, denn sie könnten nicht selbsterklärend Personen, Orte und Zeitverläufe identifizieren. Sprachtexte hingegen haben das Problem, dass sie weniger glaubwürdig sind, denn man kann vieles sagen, ohne den Beweis für die Wahrheit des Gesagten wirklich antreten zu können. Trotz allem, was wir über die Manipulierbarkeit auch von fotografischen Bildern wissen, glauben wir Bildern mehr; darüber hinaus geben sie uns eine rasche und plastische Anschauung von Ereignissen. Wir „machen uns ein Bild“, das uns rasch orientiert und universell verständlich erscheint, jenseits aller sprachlicher Besonderheiten.

Mit Bildsemantisierung und Bildaufschließung einerseits, Sprachveranschaulichung und Sprachauthentisierung andererseits sind also die sehr elementaren wechselseitigen Sprach-Bild-Bezüge umrissen. Dazu kommen aber noch eine Reihe subtilerer Leistungen von Sprache und Bild, die im Sinne kultureller Praktiken arbeiten und die ich noch kurz andeuten möchte.

### 3.2 Kulturell verankerte Praktiken

Wie Sprache und Bild kombiniert werden, hängt nicht nur von den medialen und technischen Bedingungen ab, welche die Kommunikationsform für die jeweilige Textsorte vorgibt, es bilden sich auch kulturelle Muster dafür aus, wie man im Zusammenspiel der Zeichenarten nach bestimmten Erwartungen und Normen verfährt. In unserem Beispiel geht es um den durchaus brisanten Fall einer Erschießung durch die Polizei „auf offener Straße“. Dies berührt medienethische und politische Aspekte, die auch bei der Frage, was jeweils sagbar und zeigbar ist, ins Spiel kommen.

Zum ersten medienethischen Aspekt kann festgestellt werden, dass reale Bilder, also fotografische Bilder von Sterbenden in unserer Kultur nicht gezeigt werden können, sehr wohl aber kann darüber gesprochen werden. Entsprechend kann der Sprechertext in unserem Beispiel in Satz (II) lauten: *die beamten erschossen den mutmaßlichen attentäter nach einer kurzen verfolgungsjagd auf offener strasse*. Währenddessen sehen wir die Einstellungen (2) bis (4), die Bilder von der Verfolgungsjagd (2) und das Ergebnis der Erschießung (4) zeigen; dazwischen ist eine Aufnahme einer Absperrkette von Polizisten (3) montiert, die erklärt, warum wir nicht den ganzen Ereignishergang sehen.



(3) kurzen verfolgungsjagd auf offener



(4) straße (III) er soll zuvor versucht haben  
in das regierungsgebäude

Für die „Bildlücke“ gibt es also eine ethische, aber auch eine praktische Erklärung: Das Bild von der Erschießung hat es vielleicht nicht gegeben, weil der Kameramann nicht filmen konnte, aber man hätte es auch aus Gründen der Pietät nicht zeigen können. Die Sprache kann hier einspringen, die Bilder „um den Ereignis-Kern herum“ fungieren authentisierend.

Zwar ist der Sprachtext in der Lage, sehr differenziert zu referieren und zu präzisieren, gleichzeitig bleibt der Geltungsanspruch der Aussagen aber immer prekär. Deshalb enthält er immer wieder nachrichtentypische Hecken-  
ausdrücke, die das Gesagte immunisierend relativieren: (I) *offenbar*, (II) *mut-*



*maßlichen*, (III) *soll*, (IV) *angeblich*. Zugleich werden sprachlich mehrere Bewertungen vorgenommen, durch deontische Wörter wie *bombenanschlag*, *attentäter*, *sprengsatz*, *linksextremistisch*, *terrorgruppe*, die allesamt moralische und sehr starke Negativ-Beurteilungen des Erschossenen zum Ausdruck bringen oder implizieren. Dagegen kann man der Formulierung *erschossen ... auf offener straße* die Charakterisierung des polizeilichen Vorgehens im Sinne eines Skandals entnehmen. Schließlich ist in einem Rechtsstaat äußerst strittig, ob und unter welchen Umständen die Polizei einen Verdächtigen durch Schüsse mehr als außer Gefecht setzen darf. Insofern ist eine „Erschießung auf offener Straße“ durch die Polizei immer eine höchst brisante Angelegenheit, was hier allerdings nur sehr zurückhaltend angedeutet wird.

Die unterschiedlich deutlichen und unterschiedlich gradierten Bewertungen des Erschossenen und der Polizei können auch als ausgleichende Kommentare zu den Bildern gelesen werden, die umgekehrt zwar das gewaltsame Vorgehen der Polizei zeigen, nicht aber das des mutmaßlichen Attentäters, der sichtbar nur als Opfer in Erscheinung tritt, nicht als Täter. Diese hinzugefügten Kommentare des Sprechtextes erfüllen eine „Relais“-Funktion – so charakterisiert Barthes (1990, S. 43ff.) sprachliche Bedeutungen, „die im Bild nicht aufscheinen“ (ebd., S. 36). Anders als die sprachliche ‘Verankerung’, die nur die Funktion hat, aus dem polysemen Angebot des Bildes auszuwählen und es damit durch Selektion zu kontrollieren, gewissermaßen wie mit einem Anker festzumachen, liefert die Sprache in der Relaisfunktion einen eigenen Beitrag zum intermedialen Textganzen. Insofern transkribieren die wahrheitswertabschwächenden und abgestuft evaluativen Sprachelemente die Bilder in Richtung einer Relativierung und einer entdramatisierenden Tendenz, die letztlich den gesamten Vorfall, vor allem das Vorgehen der Polizei „normalisiert“, d.h. als irgendwie balanciert einordnet. Auch die Qualität der Einstellungen und des O-Tons kann in diese Richtung wirken. So geben die Einstellungen (1) und (2) durch die Bewegung einen unmittelbaren Eindruck vom aktuellen Ablauf der Szene, die Kamera fängt mit wackelnden Bildern direkt ein, was als Ereignis berichtenswert ist, den Fluchtversuch des Protagonisten und die Verfolgung durch die Polizei, (2) enthält dazu noch Tonaufzeichnungen der Schüsse, und das ‘Ergebnisbild’ (4) zeugt durch zwei Elemente von der Dramatik der Szene: Es ist erkennbar mit einem Teleobjektiv durch die Absperrungskette der Polizei gefilmt, zeigt also etwas, was nicht gezeigt werden soll; und man sieht den Bombenentschärfer in seinem gespenstisch wirkenden Schutzanzug, was als Anzeichen für die situative Gefahr gedeutet werden kann.

Zusammenfassend kann man die wechselseitige Sprach-Bild-Transkription des kurzen ZDF-Berichts als ein wohlbalanciertes Zusammenspiel von bild-



lich vermitteltem Plot, der – trotz der ethisch und praktisch begründeten Auslassung der Todesbilder – durchaus eine gewisse Dramatik vermittelt, und einem im Meldungsstil sachlich verankernden Sprachtext sehen, der durch detaillierte Propositionen die Bilder aufschließt, dabei bewertend kommentiert und mit relativierenden bzw. legitimierenden Formulierungen alles Brisante wieder normalisiert.

In dieser Kombination zeigen sich auch kulturelle Verfahren, das Zusammenspiel von Sprache und Bildern nach politischen und ethischen Gesichtspunkten zu arrangieren, so dass die medial bedingten Verfahren noch einmal entsprechend überformt werden. Eine medienlinguistische Analyse muss deshalb beide Schichten herausarbeiten, die technisch-mediale, von der Kommunikationsform bestimmte, und die kulturelle, die Erfordernisse der Textsorte und des gesamten Kontexts zu berücksichtigen hat.

#### **4. Literatur**

Barthes, Roland (1990): *Der entgegenkommende und der stumpfe Sinn. Kritische Essays III*. Frankfurt a.M.

Holly, Werner (2009): *Der Wort-Bild-Reißverschluss. Über die performative Dynamik der audiovisuellen Transkriptivität*. In: Feilke, Helmuth/Linke, Angelika (Hg.): *Oberfläche und Performanz. Untersuchungen zur Sprache als dynamischer Gestalt*. (= Reihe Germanistische Linguistik 283). Tübingen, S. 389-406.

Holly, Werner (2011): *Bildüberschreibungen. Wie Sprechtexte Nachrichtenfilme lesbar machen (und umgekehrt)*. In: Diekmannshenke, Hajo/Klemm, Michael/Stöckl, Hartmut (Hg.): *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. (= Philologische Studien und Quellen 228). Berlin, S. 233-253.